

Exhibit A

[Submitting Counsel below]

**UNITED STATES DISTRICT COURT
NORTHERN DISTRICT OF CALIFORNIA
SAN FRANCISCO DIVISION**

IN RE: UBER TECHNOLOGIES, INC.
PASSENGER SEXUAL ASSAULT
LITIGATION

MDL No. 3084 CRB

**DECLARATION OF DOUGLAS
FORREST IN SUPPORT OF
PLAINTIFFS' POSITION FOR THE
PARTIES JOINT DISCOVERY LETTER**

This Document Relates to:
ALL ACTIONS

DECLARATION OF DOUGLAS FORREST

I, Douglas Forrest, state and declare as follows:

1. I am the Senior Vice President, eDiscovery Analytics & Strategy, at International Litigation Services ("ILS"), which is located in Irvine, California (www.ilsteam.com). I have been retained as a consultant for the Plaintiffs in this action for my knowledge relating to Electronically Stored Information ("ESI") as it relates to creating an ESI Order. The facts stated in this Declaration, except as otherwise explicitly noted, are within my own personal knowledge and, if called as a witness to testify, I could and would competently testify to the facts contained in this Declaration.

2. This is my second declaration in this matter. The first was the Declaration of Douglas Forrest (February 12, 2024), submitted as Exhibit 6 to Plaintiffs' Brief in Support of Proposed ESI Order (Case 3:23-md-03084-CRB, Document 261, Filed 02/12/24).

1 3. I make this Declaration in support of Plaintiffs' Position in the Joint Discovery
2 Letter, dated April 12, 2024 ("Letter"). I am familiar with the issues raised in the Letter and have
3 participated in several meet and confers with the parties. I have reviewed, *inter alia*:

- 4 a. Pretrial Order No. 9: Order on ESI Protocol Disputes (ECF No. 345) ("PTO
5 No. 9").
- 6 b. Letter, K Smith to R Luhana et. al, dated January 16, 2024 ("Uber Data
7 Sources Letter").
- 8 c. Letter, K Smith to R. Luhana et. al, dated February 22, 2024.
- 9 d. Plaintiffs' Brief in Support of Proposed ESI Order (ECF No. 261), dated
10 February 12, 2024.
- 11 e. Plaintiffs Proposed ESI Order (ECF No. 261-2), dated February 12, 2024.
- 12 f. Defendants' Brief in Support of Proposed ESI Order (ECF No. 262), dated
13 February 12, 2024.
- 14 g. Defendants' Proposed ESI Order (ECF No. 262-2), dated February 12,
15 2024.
- 16 h. Declaration of Phillip Favro ("Favro Decl.") (ECF No. 262-8), dated
17 February 12, 2024.
- 18 i. Declaration of Jason Alsobrook ("Alsobrook Declaration") (ECF No. 262-
19 9), dated February 12, 2024.
- 20 j. Declaration of Sam Yang ("Yang Decl."), *In re: Meta Pixel Healthcare*
21 *Litigation*, No. 22-cv-0358-WHO (VKD), ECF. No. 265 (N.D. Cal. June 1,
22 2023)
- 23 k. Declaration of Jamie Brown ("Brown Decl."), *In Re: Meta Pixel Healthcare*
24 *Litigation*, No. 22-cv-0358-WHO (VKD), ECF No. 266 (N.D. Cal. June 1,
25 2023).
- 26 l. Email, J Wikler to M Sweet et. al., dated April 9, 2024, at 9:44:38 PM EDT
27 ("Wikler Email").

28 **I. QUALIFICATIONS**

29 4. I am a graduate of Stanford Law School, where I was a Note Editor of the Law
30 Review. I was admitted to the bar in 1977 (I currently have retired status), and, after practicing law
31 at Breed, Abbott & Morgan and Cravath, Swaine & Moore, I developed expertise in computer
32
33
34
35

1 technology and software design, programming, and implementation, both generally and with
2 respect to litigation support and e-discovery.

3 5. As an attorney at Cravath, I relied on *Aquarius*, the first large-scale implementation
4 of computerized litigation support, which was implemented on the IBM antitrust cases.

5 6. As Director of Litigation Services at Legal Information Technology, Inc. (“LIT”), I
6 was instrumental in introducing imaging, coding, and search technology for discovery to Am Law
7 200 law firms, and I pioneered the practice of integrating imaging with legacy search systems such
8 as BRS.

9 7. As a systems architect, application designer and programmer, I created case
10 management, litigation support and document repository systems and applications (including
11 *WIDE*, and *LIT CaseWorks for Lotus Notes*, developed for LIT), SaaS (Software as a Service)
12 knowledge management applications (including *LexisNexis Total Alerts* and *LexisNexis Clipper*
13 developed for Ozmosys (<https://www.ozmosys.com/>)), and e-discovery and production operation
14 systems for Doar (<https://www.doar.com/>). The software that I designed, programmed, and
15 implemented to produce *LexisNexis Total Alerts*, *LexisNexis Clipper*, *LexisNexis Legal Industry*
16 *Monitor*, *Thompson Elite Daily Docket*, and *Institutional Investors’ Mutual Fund Daily*, *Hedge*
17 *Fund Daily* and *Compliance Daily*, traversed thousands of web sites and extracted thousands of
18 URLs from those sites daily.

19 8. At ILS, I direct e-discovery analytics and strategy, providing technical advice,
20 expertise and drafting and other assistance to counsel.

21 9. I have advised, consulted, or acted as a declarant or affiant with respect to ESI in
22 many cases, including:

23 a. *In re: Social Media Adolescent Addiction/Personal Injury Products Liability*
24 *Litigation*, MDL No. 3047 (N.D. Cal.);

25 b. *In Re: StubHub Refund Litigation*, Case No. 4:20-md-02951-HSG (N.D. Cal.);

c. *In Re: Meta Pixel Healthcare Litigation*, Case No. 3:22-cv-3580-WHO (N.D. Cal.);

- d. *In Re: Philips Recalled CPAP, Bi-Level Pap, and Mechanical Ventilator Products Litigation*, MDL No. 3014 (W.D. Pa.);
- e. *Nichols v. Noom Inc.*, No. 20-CV-3677 (LGS) (KHP), 2021 WL 948646, (S.D.N.Y.);
- f. *In re: Ethiopian Airlines Flight ET 302*, Lead Case: 1:19-cv-02170 (N.D. Ill.) (Boeing 737 Max Crashes);
- g. *In Re: 3M Combat Arms Earplug Product Liability Litigation*, MDL No. 2885 (N.D. Fla.) (“3M Combat Arms MDL”);
- h. *In Re: Volkswagen “Clean Diesel” Marketing, Sales Practices and Products Liability Litigation*, MDL No. 2672 (N.D. Cal.) (“VW Clean Diesel MDL”);
- i. *In Re: Intel Corp, CPU Marketing, Sales Practices and Products Liability Litigation*, MDL No. 2828 (D. Or.);
- j. *In Re: Chrysler-Dodge-Jeep EcoDiesel Marketing, Sales Practices and Products Liability Litigation*, MDL No. 2777 (N.D. Cal.);
- k. *Lafferty v. Alex Jones* (Conn. Super. Ct.);
- l. *Soto v. Bushmaster Firearms International* (Conn. Super. Ct.) (Sandy Hook parents);
- m. *In Re: Takata Airbags Product Liability Litigation*, MDL No. 2599 (S.D. Fla.);
- n. *In Re: Testosterone Replacement Therapy Products Liability Litigation*, MDL No. 2545 (N.D. Ill.);
- o. *In Re JCCP 4771, Zolofit Birth Defect Cases*, (Cal. Super. Ct.);
- p. *Da Silva Moore v. Publicis Groupe & MSL Group*, No. 11 Civ. 1279 (ALC)(AJP) (S.D.N.Y.) (seminal TAR case).

10. I have served as a speaker or panelist on many CLE webinars including *Balancing the Needs of Requesting and Producing Parties: Getting E-Discovery Right* (RAND Institute for Civil Justice Conference, October 3-4, 2023) (“RAND Conference”), *An Analysis of Today’s Mass Tort Landscape Agenda* (HarrisMartin MDL Conference, March 27, 2019), *Current Mass Torts from E-Discovery Through Exit Strategies – Navigating “Game-Changing” Dynamics* (HarrisMartin MDL Conference, November 26, 2018), *The Mass Tort Litigation Landscape – A Critical Analysis Agenda* (HarrisMartin MDL Conference, September 26, 2018), *The State of E-*

1 *Discovery in 2018: Analysis & Review* (West LegalEdCenter, September 27, 2018), *Lessons*
2 *Learned from Recent eDiscovery Disasters* (West LegalEdCenter, February 26, 2018), and *Top*
3 *ESI Mistakes Made in Mass Tort Disputes* (West LegalEdCenter, September 14, 2017).

4 11. Along with Uber's expert Dr. Maura Grossman, I was an advisor on and assisted in
5 setting up the RAND Conference.

6 12. I am a member of the drafting team for forthcoming commentary of the Sedona
7 Conference Working Group 1 (<https://thesedonaconference.org/wgs/wg1>) on *Conducting*
8 *eDiscovery of Modern Communication and Collaboration* and was on the panel presenting the
9 *Drafting Team Report: Discovery of Modern Communications and Collaboration Platforms* at the
10 Sedona Conference Working Group 1 Annual Meeting in October 2023.

11 **II. OPINIONS**

12 **A. *Questions Addressed***

13 13. In PTO No. 9, the Court stated:

14 “[I]n recognition of the challenging nature of hyperlinks, Uber shall direct an employee
15 with knowledge and expertise regarding Google Vault and Uber’s data and information
16 systems to investigate in detail the extent to which Google Vault’s API, macro readers,
17 Metaspike’s FEC or other programs may be useful to automate, to some extent, the process
18 of collecting the contemporaneous version of the document linked to a Gmail or other
19 communication within Uber’s systems, whether the email or communication is stored in
20 Google Vault, or outside. This investigation shall not be limited to documents referenced
21 by URL or hyperlinks in emails or Google documents stored in Google Vault, but shall
22 also include other cloud-based messages such as Slack. Uber’s designated employee may
23 consult with Uber’s e-discovery experts. Likewise, Plaintiffs shall also more thoroughly
24 investigate these potential solutions.”¹

24 ¹ ECF No. 345 at 21.

1 14. In this Declaration, I describe the steps undertaken for Plaintiffs by ILS under my
2 direction to investigate potential solutions and the conclusions that I reached with respect to them.

3 B. *Background Circumstances*

4 15. My understanding is that:

- 5 a. Uber uses Google Workspace, including native Google documents² in
- 6 Google Drive, Google Mail and Google Vault.
- 7 b. Uber deletes Google Mail not subject to a legal hold after six months.
- 8 c. Uber's Google Drive data has been, and continues to be, preserved
- 9 indefinitely.
- 10 d. Uber preserves at least some Google Drive documents and Google Mail
- 11 with Google Vault.

12 16. When a Google Vault document is exported from Google Vault, the current version
13 of that document is exported as the corresponding Microsoft document type, e.g., Google Docs
14 documents are exported as Microsoft Word documents.

15 17. Uber's vendor Lighthouse has received separate collections of Google email and
16 Google Drive documents exported from Google Vault. These exported documents are not the
17 native Google documents, with links to prior revisions, and Google mail as they exist within
18 Google Workspace but rather versions of those documents converted into the corresponding
19 Microsoft file types. These documents are then processed and loaded into *Relativity*.

20 18. Lighthouse has a Google Parser which can identify hyperlinks which include
21 Google DocumentIDs. Lighthouse then runs a script to determine if documents with those
22 DocumentIDs are in the collection of exported documents and creates metadata to associate any
23 documents found with their parent Google emails.

24 C. *The Parties' Positions*

25 ² Google Docs, Google Sheets, etc. Unless otherwise noted, all references to Google Drive documents in this
declaration refer to Google native file types.

1 19. Uber's Google email exists in two buckets. Bucket one is emails preserved in
2 Google Vault ("Google Vault Email") which, depending on age, may or may not still exist in
3 Uber's active Google email environment. Bucket two is undeleted Google emails in Uber's
4 workspace, i.e., emails sent or received within the last six months and going forward which have
5 not been deleted under Uber's retention policy ("Active Google Email").

6 20. The linked documents also exist in two buckets, Bucket one is Google Drive
7 documents preserved in Google Vault ("Google Vault Documents"). Bucket two is undeleted
8 documents in Google Drive ("Google Drive Documents").

9 21. The parties do not have a disagreement with respect to instances where the current
10 versions of documents as exported to Lighthouse are in fact the most recent versions extant when
11 an email was sent. i.e., when the date of an exported (current) version is before the date of its
12 parent email.

13 22. The Google Vault Search interface provides a way to manually retrieve, if it exists,
14 of the most recent version of a document dated before the date of an email with a link to it.
15 Plaintiffs' understanding is that Uber has proposed using this manual method but has not
16 specified, in writing or otherwise, precisely what those steps are or how long it would take
17 someone experienced in their process to retrieve a contemporaneous version.

18 23. Plaintiffs propose a solution constituted by three components.

19 24. The first component, for Active Google Email is using Metasploit's *Forensic Email*
20 *Collector* ("FEC"), a program which can retrieve active Google email and the contemporaneous
21 versions of linked Google Drive documents. FEC is a proven program which has been used for
22 years by organizations and governmental entities such as Paul Weiss, PriceWaterhouseCoopers,
23 the Office of the California Attorney General, FTI Consulting, Deloitte, Grant Thornton, Clifford
24 Chance, Stroz Friedberg, Winston & Strawn, the Federal Trade Commission, the Norwegian Tax
25 Administration, and the Australian Federal Police.

 25. The second component addresses links to Google Drive documents in Google
emails that have been deleted from Google email and thus cannot be retrieved using FEC.

1 Plaintiffs propose that these contemporaneous documents be retrieved programmatically using the
2 Google Drive API. Plaintiffs have created a proof of concept program (“POC”), described in more
3 detail at Paragraphs 29-38 below which demonstrates that such a program can be written. ILS is
4 continuing the development of this program to make it production ready, again just to demonstrate
5 that a production ready version could be developed quickly.

6 26. The third component, for any documents which cannot be retrieved
7 programmatically, either using FEC (to the extent that the linking emails are still active) or the ILS
8 program, is to use Uber’s manual method (once it has been sufficiently described in writing). The
9 number of documents to be so retrieved should be significantly and substantially reduced, as it
10 will not include any documents retrieved by FEC or the production ready version of the ILS
11 program.

12 D. *Plaintiffs’ Thorough Investigation*

13 27. Following the issuance of PTO No. 9, Plaintiffs thoroughly investigated “in detail
14 the extent to which Google Vault’s API, macro readers, Metaspike’s FEC or other programs may
15 be useful to automate, to some extent, the process of collecting the contemporaneous version of
16 the document linked to a Gmail or other communication within Uber’s systems.”

17 28. Plaintiffs’ investigation proceeded on two tracks. One track was attempting,
18 through multiple meet and confers and written questions to Uber, to find out what Uber had
19 attempted or discovered through their investigation. I participated in both meet and confers with
20 Uber on March 27 and April 2, 2024. Another track was ILS’s investigation into the Google Vault
21 API and macro recorders as well as other programs such as the Google Drive API which could be
22 useful to automate the process.

23 a. *ILS’s Program to Collect Contemporaneous Versions of Google Drive Documents*
24 *Using the Google Drive API*

25 29. Despite Uber’s adamant position that it was not possible to create a program that
collected the contemporaneous versions of Google Drive documents from links in their parent

1 emails, I knew that it was possible to create such a program because Metaspike had already done it
2 with FEC.

3 30. However, FEC works only on live Gmail collected by FEC, which was not
4 possible here with respect to Gmail which had already been deleted and existed, if at all, only in
5 Google Vault.

6 31. I reasoned that Metaspike must be using the Google Drive API to retrieve the
7 correct contemporaneous Google Drive documents contained in the parent Google emails that it
8 processed.

9 32. Therefore, I knew that ILS could also use the Google Drive API for the same
10 purpose but using Google Drive link data that had been extracted from Google Vault exports
11 instead of from live emails. What I didn't yet know, despite extensive research, was how to do
12 this.

13 33. On the morning of April 2, 2024, the date of Plaintiffs' last meet and confer with
14 Uber, I found a post³ in Stack Overflow ("Stack Overflow post"), a well-known and widely used
15 forum for developers, that set out a method to programmatically retrieve a date-specific revision
16 of a Google Drive Google native document identified by its DocumentID.

17 34. Later that day, using the methodology set out in the code in the Stack Overflow
18 post that I had found and working under my direction, , an ILS programmer created a Proof of
19 Concept program ("POC") which demonstrated that the following operations could be executed
20 programmatically without manual intervention⁴:

- 21 a. Loading service account credentials;
- 22 b. Impersonating a user based on their email address and generating
23 appropriate access credentials with the service account;
- 24

25 ³[https://stackoverflow.com/questions/77467875/revert-to-specific-version-of-google-sheets-with-respect-to-specific-date-using#:~:text=This%20sample%20script%20is%20for%20Drive%20API%20v3,%60No%20revisions%20in%20%24%7Bdate%7D.%60\)%3B%0A%20%207D%0A%7D](https://stackoverflow.com/questions/77467875/revert-to-specific-version-of-google-sheets-with-respect-to-specific-date-using#:~:text=This%20sample%20script%20is%20for%20Drive%20API%20v3,%60No%20revisions%20in%20%24%7Bdate%7D.%60)%3B%0A%20%207D%0A%7D)

⁴ A one-time log in into the Google API website is required to acquire service account credentials.

- c. Using those credentials, retrieving the revision list of a specific Google Drive document identified by its DocumentID;
- d. Iterating through the revision list and identifying the most recent revision before a specific date;
- e. Downloading that revision as its corresponding Microsoft file type.

35. Development and testing of this program was completed on April 3, 2024.

36. Based on this development and testing, I drafted Plaintiffs Proposed Methodology for Retrieving Google Drive Documents Linked to Within Google Emails (“Plaintiffs Proposed Methodology”), which set forth a complete step-by-step methodology by which Uber could do just that, based upon the Stack Overflow post methodology referenced in the proposal. It is my opinion that Plaintiffs Proposed Methodology is viable, feasible, and not unduly burdensome. The steps are set forth below:

- a. First, Uber’s vendor, Lighthouse, would extract the links to Google Drive documents from Google Mail using Google Parser which does this extraction while preserving the metadata establishing the relationship of each extracted link to its parent Google email.
- b. Second, using the extracted links identified supra, Lighthouse would identify: (1) the Google Drive documents in its possession which correspond to those links and (2) the links for which Lighthouse does not have a corresponding Google Drive document.
- c. Third, if Lighthouse has a corresponding linked Google Drive document, Lighthouse would compare the date and time of the linked Google Drive document against the sent date and time of the Gmail using Google parser or SQL queries in the SQL database. If the document date and time is less than or equal to the Gmail sent date and time, then the Google Drive

document is the correct contemporaneous one and Lighthouse will provide metadata that will enable the documents to be linked as family groups for production.

- d. Fourth, if Lighthouse does not have the linked Google Drive document in its possession, or if the copy is not contemporaneous, it would be added to a “Missing List” which should include metadata fields including: (1) the ID (BegBates) of the Gmail which contained the link, (2) the Sent Date and time of that email, (3) the sender of the email, including email address, (4) the recipients of the email, including email addresses, and (5) the link URL for the Google Drive document.
- e. Fifth, a programmer with Google Drive API expertise and appropriate access to Uber’s Google Workspace environment, will, using the Stack Oversight post code as a foundation, create a program that cycles through the Missing List and for each link extracts the Google DocumentID and then retrieves its revision list recording whether the retrieval was successful or failed, with any error codes if it failed. Next, if the retrieval of the revision list was successful, cycle through the revision list and compare the listed revisions from latest to earliest until the latest listed revision preceding the parent email’s sent date and time is identified as the revision of the correct contemporaneous document. If no such revision can be located, record that failure and proceed to process the next link in the Missing List. Lastly, if the correct revision is located, convert the file to its Microsoft equivalent and download it along with appropriate metadata. If that operation fails, record that failure, and proceed to process the next link in the Missing List.
- f. Sixth, Lighthouse would link each retrieved document with its metadata to its parent email for production.

1 g. Seventh, Lighthouse will produce responsive non-privileged emails along
2 with corresponding Google Drive documents and metadata linking the
3 documents to its parent email. Metadata for these documents will include
4 an additional “Non-Contemporaneous” metadata field which will contain
5 the value “Y” if a document was not the correct contemporaneous
6 document.

7 h. Lastly, to record the production status of missing family members, the
8 produced metadata for the parent email identified supra will include an
9 additional metadata field “Missing Google Drive Attachments” which will
10 include the links to all Google Drive documents which could not be
11 retrieved.

12 37. Summing up, ILS’ POC demonstrates that it is possible to do exactly what is
13 necessary here: to programmatically retrieve the versions of linked Google Drive documents that
14 are contemporaneous with the Google emails that contain those links.

15 38. ILS has continued to develop this program to make it production ready, again just
16 to demonstrate that a production ready version could be developed quickly.

17 *b. Defendant’s Objections to the ILS Program Are Groundless*

18 39. Uber’s criticisms and objections to Plaintiffs Proposed Methodology are
19 groundless.

20 40. First, as discussed above in paragraphs 29-37, ILS’s POC demonstrates that such a
21 program can be written.

22 41. Second, with respect to document links in undeleted emails, there already is such a
23 program, viz., FEC as discussed in paragraphs 72-76 *infra*.

24 42. Third, ILS is continuing the development of this program to make it production
25 ready, again just to demonstrate that a production-ready version could be developed quickly.

43. Fourth, both the POC and the production ready version handle Uber’s enumerated
failure points through standard error capture and logging without manual intervention.

1 44. Fifth, capturing revision id data, which the program would have at the time of
2 retrieval, as metadata fields which would be included in a load ready production is a trivial task.
3 This is basically the same sort of process as Lighthouse does with the Google DocumentIDs that it
4 extracts with its Google Parser. *See* Alsobrook Declaration, Paragraphs 10-11.

5
6 *c. The ILS Contemporaneous Google Drive Document Program Will Work With Other*
7 *Types of Parent ESI*

8 45. PTO No. 9 states that Uber’s “investigation shall not be limited to documents
9 referenced by URL or hyperlinks in emails or Google documents stored in Google Vault, but shall
also include other cloud-based messages such as Slack.”⁶

10 46. My understanding is that Defendant has not proposed any methodology for
11 extracting links to Google Drive documents from Slack channels or direct messages, or for
12 preserving the relationship of those links to their parent Slack messages, or for retrieving any
versions, current or contemporaneous, of the Google Drive documents so linked.

13 47. Plaintiffs’ proposal for retrieving Google Drive documents from Gmail would also
14 work for links to documents in Uber’s Google Drive extracted from Slack channels and direct
15 messages where the author’s email address is available.

16 48. The Uber Data Sources Letter also lists additional platforms, systems, and
17 applications such as Google Chat, uChat⁷, and HipChat messaging systems, Bliss and Zendesk
18 customer support systems, the PureCloud and LiveOps phone systems, as well as the JIRA
19 ticketing system.

20 49. It is my understanding that some of these additional platforms, systems and
21 applications may also include links to Uber Google Drive documents.

22 50. It is also my understanding that Uber has not proposed any methodology for
23 extracting links to Google Drive documents from these additional platforms, systems, and

24 ⁶ ECF No. 345 at 21.

25 ⁷ uChat is an internally developed Uber messaging application that was in use from February 2017 to 2020. Uber Data Source Letter, p. 6.

1 applications, or for preserving the relationship of those links to their parents, or for retrieving any
2 versions, current or contemporaneous, of the Google Drive documents so linked.

3 51. Plaintiffs' proposal for retrieving Google Drive documents from Gmail would also
4 work for links to documents in Uber's Google Drive in any of these Uber platforms, systems, and
5 applications where the author's email address is available.

6 *d. Uber's Use of Links to Google Drive Documents*

7 52. It is my understanding that Uber's vendor *Lighthouse* has received and processed
8 separate exports of unspecified collections of Gmail and Google Drive documents from Google
9 Vault and loaded the processed ESI, including extracted text, into its litigation support platform
10 *Relativity* and its underlying SQL database.

11 53. *Lighthouse*'s Google Parser can extract links to Google Drive documents from
12 Gmail. It is my opinion that there is no reason that *Lighthouse* could not easily⁸ extract all other
13 links, e.g., to documents in Box, from Gmail as well.

14 54. It is my understanding that Uber has thus far refused to disclose how many
15 exported Gmails and Google Drive documents have been provided to *Lighthouse* and what
16 *Lighthouse* and Uber know about the prevalence and composition of the links to Google Drive and
17 other document repositories in those exports.

18 55. The Wikler Email provides some estimates of the prevalence of Google email links
19 to Google Drive documents in "tens of millions of documents":

20 *f.* 7% of emails have links to Google Drive documents/.

21 *g.* There were an average of 2 links to Google Drive documents in Google
22 emails that contained links.

23 *h.* 20% of the Google Drive documents were contemporaneous with the Gmail
24 that linked to them.

25 ⁸ As noted in paragraph 9 *supra*, I designed, programmed, and implemented the software to produce LexisNexis Total Alerts, LexisNexis Clipper, LexisNexis Legal Industry Monitor, Thompson Elite Daily Docket, and Institutional Investors' Mutual Fund Daily, Hedge Fund Daily and Compliance Daily. This software traversed thousands of web sites and extracted thousands of URLs from those sites daily. Link extraction is a common task, and not a difficult one.

1 56. The flip side of the 20% estimate of contemporaneous documents is that 80% of the
2 linked Google Drive documents were not contemporaneous with the Gmail that linked to them.

3 57. Doing the math on a per million documents basis works out to 70,000 emails with
4 links to Google Drive Documents (7%), 140,000 links (70,000 x 2 links per email), 28,000
5 contemporaneous documents (20% of 140,000 links), and 80,000 (140,000 – 28,000) non-
6 cotemporaneous, post-dated documents.

7
8 58. These metrics are important in at least two ways. First, the volume of documents is
9 a consideration in proposing methods to address them. Second, even the estimated volumes make
10 plain the pervasive and integral role played by hyperlinks to Google Drive documents. The family
11 relationships established by such links will be critical evidence in tying specific Uber custodians
12 to knowledge of the linked documents and their contents.

13 *e. Uber’s Proposed Method of Retrieving Contemporaneous Versions of Google Drive*
14 *Documents Linked to Gmail*

15 59. Uber’s stated position⁹ is “Uber would have to perform a manual review using the
16 information from the specific email to locate the Google Drive document. Uber would then need
17 to conduct a manual search in the revision history of that Google Drive document to locate the
18 “last version saved before 12:00 AM on the specified date” of the specific email at issue.” Favro
19 Decl. ¶ 22.¹⁰

20 60. Uber has maintained this same position – that Uber can only manually pull
21 contemporaneous hyperlinked documents one-by-one – during the two meet and confers that I
22 attended..

23 ⁹ Uber also stated that “Plaintiffs’ request for historic ‘versions’ of documents is premised on the assumption that all
24 such documents were modified after emails referencing them were sent. ECF No. 262 at 14. Google Vault exports
only the current versions of Google Drive documents, therefore any exported version of a Google Drive document
with a last modified date before the sent date of a parent Gmail would be the correct contemporaneous version and no
further search for a historic version would be necessary.

¹⁰ ECF No. 262 at 13 (footnote omitted).

1 61. During the meet and confers, Uber's counsel was unwilling to document the
2 specific steps necessary in the manual process.

3 62. This refusal is continued in the Wikler letter which states that "We talked at length
4 about Uber's practical experiences in trying to collect particular versions of Google Drive
5 documents and the obstacles they have encountered. We also discussed in our conferrals the
6 investigation Uber has performed to identify a programmatic solution. Not only did we provide
7 this information during our call and discuss the investigations Uber has conducted, but we
8 explained why this process that requires manual intervention has not been automated."

9 63. Regardless of how long Defendant and its counsel and vendors may have talked
10 during the meet and confers, neither in their talking nor in prior writing nor in the Wikler email
11 has Defendant set out the specifics of any of this.

12 64. I note that even manual processes can be scaled up by using additional resources.
13 Consider a typical manual document review using contract attorneys who can review on average
14 400 documents a day. Parties ensure that such reviews are completed within a reasonable time
15 simply by hiring sufficient contract attorneys. This same concept would be true here.

16 *f. Uber's Investigation*

17 65. It is my understanding that Plaintiffs submitted over 20 questions to Uber seeking
18 written responses to ascertain what steps Uber took to investigate the Court's directives, and what
19 answers they were able to gather. Joint Letter, Exhibit G.

20 66. These questions would not only help Plaintiffs understand what steps Uber had
21 taken, but also help Plaintiffs in creating additional proposed solutions to Uber in order to
22 accomplish this Court's goal.

23 67. Uber was only willing to discuss the answers to some of these questions generally
24 on the meet and confers that I participated in and it is my understanding that Uber would not
25 provide anything specific in writing.

 68. It is my understanding that Uber's position with respect to retrieving linked
documents from Google Drive has not changed.

1 69. Uber never offered any solution other than their original position.

2 70. I have not seen any indication that Uber seriously investigated using FEC on
3 undeleted active email in Google mail in Uber's environment, e.g., by hiring a consultant with
4 FEC expertise.

5 *g. The Google Vault API and Macro Readers*

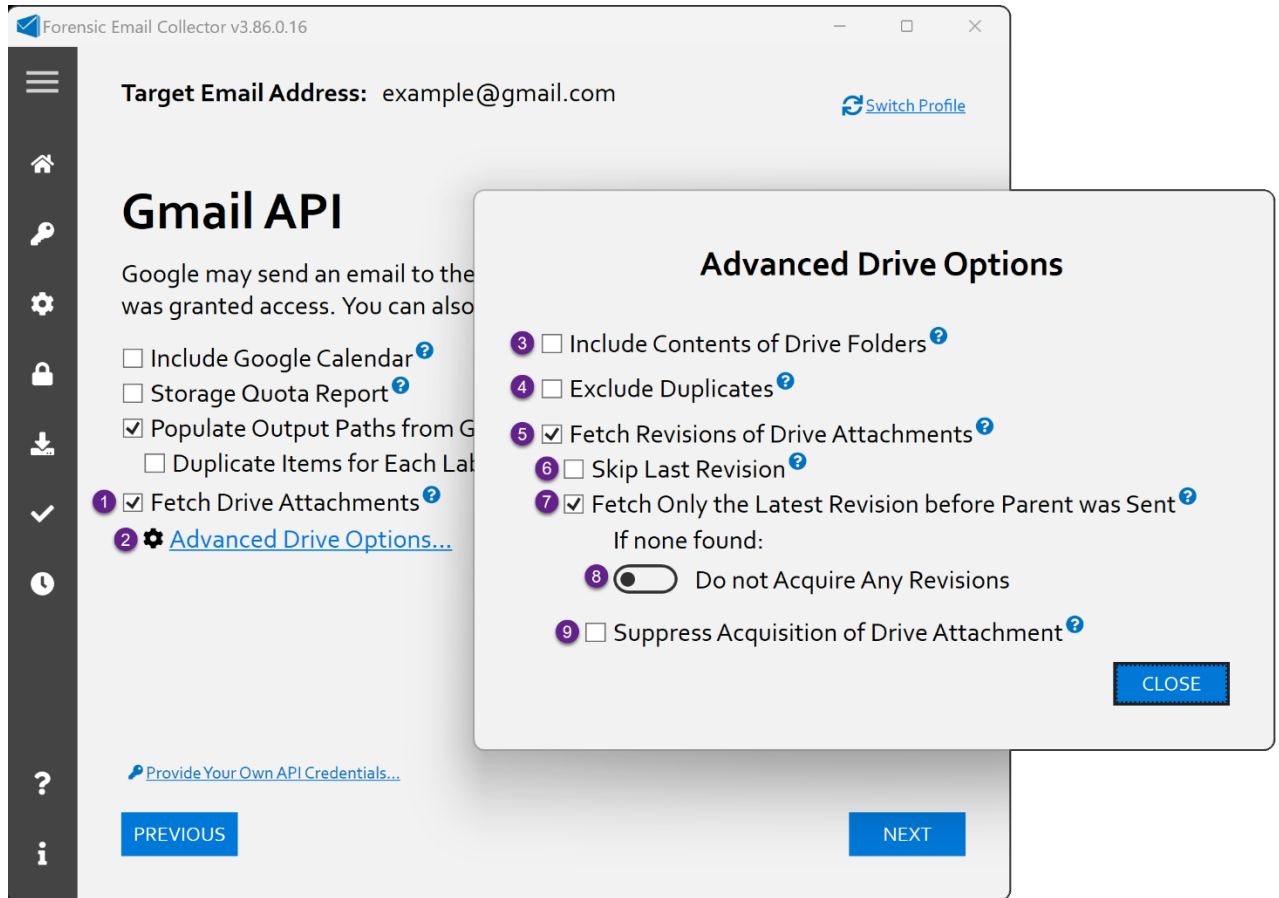
6 71. Upon further and extended investigation by ILS after entry of PTO No. 9, I
7 concluded that as of now¹¹ neither the Google Vault API nor macro recorders, which would be
8 employed within Google Vault's search interface, appeared to provide a framework for a
9 programmatic solution to retrieving the contemporaneous version of a Google Drive document.

10 *h. Metaspike's Forensic Email Collector*

11 72. Metaspike's *Forensic Email Collector* ("FEC") does provide a programmatic
12 solution for "collecting the contemporaneous version of the document linked to a Gmail"¹²:
13
14
15
16
17
18
19
20
21

22 ¹¹ Just as Google only very recently (on December 8, 2023) added the capability to collect the current versions of
23 Google Drive documents when Google emails that link to them are collected through the Google Vault search
24 interface, Google may at some point add a capability to Google Vault to collect the contemporaneous versions of such
25 documents either through the Google Vault API or the Google Vault search interface.

¹² <https://docs.metaspike.com/article/46-acquiring-google-drive-attachments-of-emails> (last accessed on April 7, 2024).



73. FEC could be used on a subset of Uber’s Google mail, viz., the portion that is still active in Uber’s Google mail environment, i.e., at a minimum the last 6 months of email which has not been deleted under Uber’s retention policy. It could also be used going forward on future Uber Google emails before Uber deletes them.

74. I discussed FEC in my earlier declaration, where I stated:

70. *Forensic Email Collector* has been used for years by demanding organizations. Prominent FEC users, including users in government and law enforcement, and in prominent firms in the Am Law 100, the accounting Big 3, and litigation support and forensic services, include Paul Weiss, PriceWaterhouseCoopers, the Office of the California Attorney General, FTI Consulting, Deloitte, Grant Thornton, Clifford Chance,

1 Stroz Friedberg, Winston & Strawn, the Federal Trade Commission, the Norwegian Tax
2 Administration, and the Australian Federal Police.¹³

3 71. Defendants have objected to the use of *Forensic Email Collector* on some
4 general grounds including that they are concerned about scale, security about using a third-
5 party vendor and did not believe it can access emails that are stored in Google Vault
6 (Uber's document retention system) and thus will be incapable of retrieving emails/linked
7 documents that are only accessible in Vault (which would be the case for older emails).

8 72. In response, I note the following:

- 9 a. *Forensic Email Collector*'s inability to access emails stored in Google
10 Vault does not affect its ability to collect and process emails and documents
11 that are *not* in Google Vault and would not have affected its ability to do the
12 same earlier when documents that now exist only in Google Vault still
13 existed outside it in Google Drive.
- 14 b. If Uber has not already tested or used *Forensic Email Collector*, any
15 security concerns that Uber has could be resolved in the same way that Uber
16 usually tests any other third-party software that Uber has introduced to run
17 on its system. Moreover, Uber has decided to use a platform, [Google
18 Workspace], which is run by third party vendors and may be subject to data
19 leaks and security breaches, versus maintaining their information in-house
20 on applications run internally on a physical server.
- 21 c. Scalability issues can be caused by many different factors, including
22 misconfiguration or sub-optimal application of a program, network
23 configuration and congestion and throttling imposed by Google.
24 Defendants' counsel have not provided any details – how many email
25 accounts, how many copies of *Forensic Email Collector* deployed, network

¹³ (Battle-Tested Software, <https://www.metaspike.com/forensic-email-collector/#customers> (last accessed on March 7, 2021)).

1 and internet configuration, etc. – about the specific deployments of Forensic
2 Email Collector with which they had issues. I note that ILS has reached out
3 to Metaspike and were informed that multiple copies of Forensic Email
4 Collector can be deployed and that it was aware of instances where Forensic
5 Email Collector was faster than Google Vault.

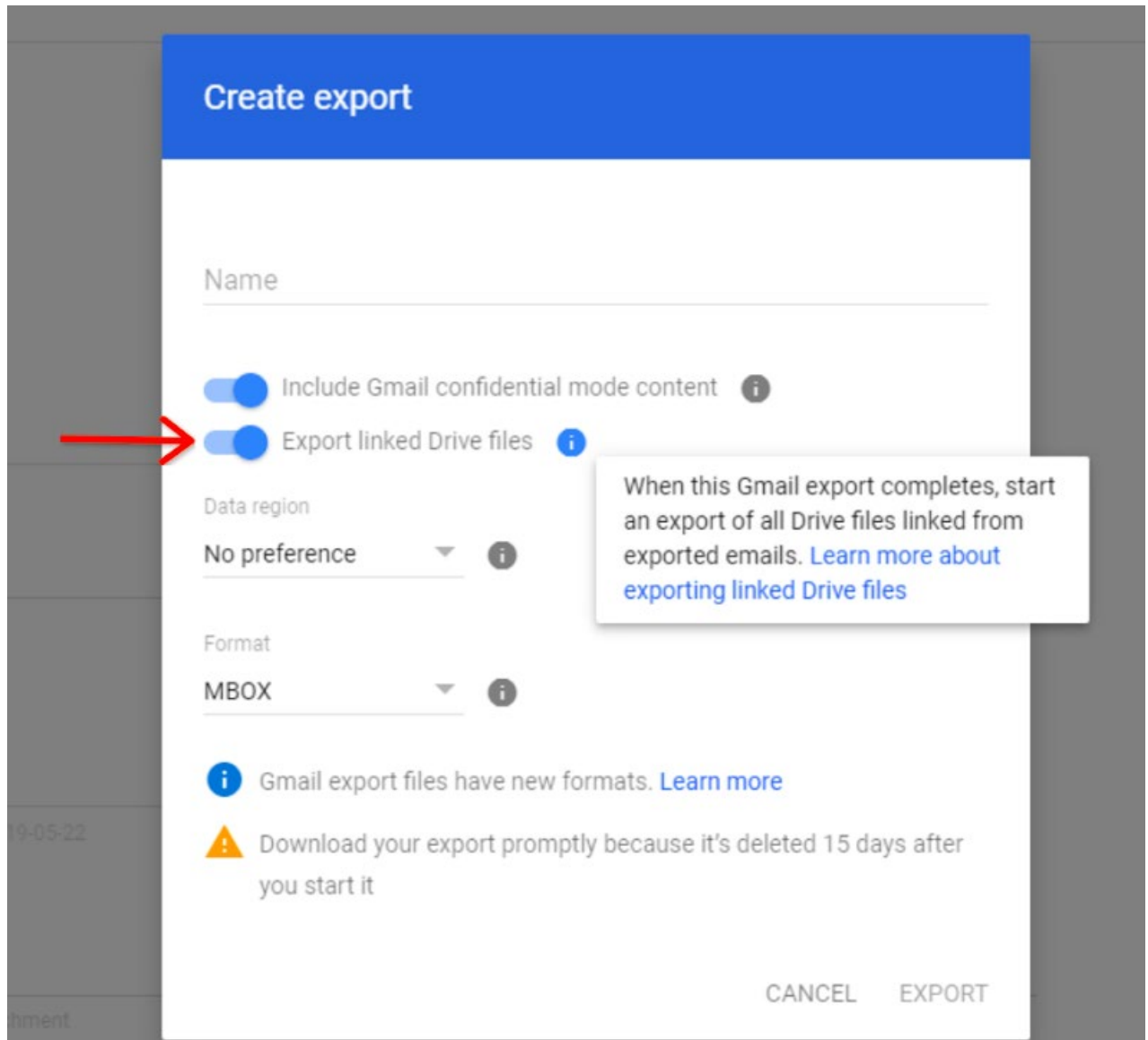
6 75. I stand by this earlier discussion and conclude that FEC is a viable programmatic
7 solution for collecting the contemporaneous versions of Google Drive documents linked to Gmail.

8 76. During the meet and confers after PTO No. 9, Uber did not offer any further
9 specific detail on its objections to FEC¹⁴, and none of Uber’s declarants cited in PTO No. 9
10 disputed that FEC could indeed collect the contemporaneous version of a document linked to a
11 Gmail.

12 *i. Google Vault vs. Lighthouse’s Google Parser*
13
14
15
16
17
18
19
20
21

22 ¹⁴ During the second meet and confer, Uber counsel admitted that Defendant’s only consultation with Google was a
23 call to Google’s general public support desk. I note that with respect to throttling by Google, Google provides specific
24 procedures for an organization to request that Google quotas which are slowing an organization’s interactions with
25 Google Workspace, i.e., throttling, be raised.
[https://support.google.com/a/answer/6301355?hl=en%20\(Last%20accessed%20April%202024#:~:text=Increase%20the%20quota,to%20be%20approved](https://support.google.com/a/answer/6301355?hl=en%20(Last%20accessed%20April%202024#:~:text=Increase%20the%20quota,to%20be%20approved). (Last accessed April 10, 2024).)

77. Since December 8, 2023, Google Vault has had the capability to retrieve the current versions of Google Drive documents attached via links in Gmail¹⁵:



78. Lighthouse's Google Parser does not have this capability as it does not access the live Google Workspace environment and does not make any use of the Google Vault API or the Google Drive API.

¹⁵ <https://workspaceupdates.googleblog.com/2023/12/google-vault-export-hyperlinked-drive-content-from-gmail-messages.html> (accessed on April 7, 2024) (cropped).

1 79. My understanding is that Lighthouse itself has had no access to Uber’s Google
2 Workspace and that its Google Parser works on exports of Gmail and Google Drive documents
3 from Google Vault.

4 80. I haven’t seen any written specifications or representations of critical details about
5 the operation of Google Parser; for example, (i) does it extract the complete URLs or just the
6 Google DocumentID segments of the Google Drive attachment links in Gmail, and (ii) what other
7 information is extracted or recorded from the parent Gmail so that the relationship of that parent
8 Gmail to a retrieved current version of Google Drive document may be preserved.

9 81. My understanding is Uber has not specified what Gmail and Google Drive
10 documents were and are in Lighthouse’s possession.

11 82. My understanding is the Google Drive exports received by Lighthouse represent
12 separate collections of as yet unspecified (i) Google Gmail and (ii) Google Drive documents. This
13 raises the possibility that there may be links in the collected Google Gmail to Google Drive
14 documents that are in Google Vault but weren’t included in the Google Vault Google Drive
15 exports being processed by Lighthouse. This issue should not arise if Google vault’s capability to
16 export linked Google Drive document when exporting parent Google emails is utilized.

17 83. My understanding from the meet and confers with Uber in which I was a
18 participant is that Uber has not seriously considered any use of this Google Vault capability,
19 which has been in production for over three months.

20 *j. Links to Documents in Box*

21 84. Box¹⁶ is a well-known and widely used cloud-based corporate document
22 repository.

23 85. The Uber Data Source Letter states that “Uber employees are also allowed to obtain
24 access to a Box account affiliated with their company email address and may use Box to store,
25 share, and collaborate on files.”¹⁷

¹⁶ <https://www.box.com>.

¹⁷ Uber Data Sources Letter, p. 6.

1 86. Uber admits that Box documents “can be placed on a legal hold” but Uber hasn’t
2 done so¹⁸.

3 87. Box has an API¹⁹, which includes functionality to retrieve prior versions of a
4 document so that contemporaneous versions can be retrieved²⁰. Some Box functionality,
5 apparently including version tracking, may be available only for Premium accounts²¹. It is my
6 understanding that Defendant has not disclosed what Box license(s) it holds.

7 88. My further understanding is that Defendant has not proposed any methodology for
8 extracting links to Box documents from Gmails or from any of Uber’s other platforms, systems, or
9 applications, or for preserving the relationship of linked Box documents to the parents that contain
10 the links, or for retrieving the linked Box documents.

10 E. *Metadata Issues*

11 89. It is my understanding, from the Wiler Email, that Uber has withdrawn its
12 objections to the inclusion of ParticipantPhoneNumbers and OwnerPhoneNumbers fields.

13 90. Plaintiffs agree to Uber dropping the Rfc822MessageID field so long as the
14 Rfc822MessageID value will be provided in the MSGID field. Wikler Email.

15 91. Uber has agreed to include the Account field, but would limit it to only ‘the e-mail
16 address associated with the single custodian whose data survives the de-duplication process.’
Wikler Email.

17 92. As so limited, email addresses associated with duplicate responsive documents,
18 would be suppressed, and Plaintiffs therefore object. A custodian may have multiple Google email
19 addresses used for different functions. Suppressing a custodian’s email address on documents
20 produced from another custodian would strip Plaintiffs of their ability to breakdown that
21 custodian’s documents by which email account it belonged to.

22
23 ¹⁸ Uber Data Sources Letter, p. 6.

¹⁹ <https://www.box.com/platform>.

²⁰ See, e.g., <https://developer.box.com/reference/get-files-id-versions-id/> (last accessed on April 8, 2024).

²¹ Ibid.

1 93. Uber has continued to object to producing the LINKGOOGLEDRIVEURLS field
2 which is defined as a “One-to-many field containing the URLs for any Google Drive linked
3 documents referenced within a given Gmail document. If a Google Drive linked document has
4 been produced, the URL will be suffixed by a space and the corresponding BEGBATES in square
5 brackets ([]).”

6 94. As Lighthouse’s Google Parser already extracts those Google Drive URLs,
7 Defendant’s burden issue would appear to be solely with suffixing the URLs with the Bates
8 numbers forof those Google Drive documents which have been produced.

9 95. The reason for adding the suffixed Bates numbers was to identify which Google
10 Drive documents, if any and for whatever reason, had not been produced. The URLs of these
11 unproduced documents would be identifiable because they wouldn’t have a suffixed Bates numbe,
12 an absence which would also establish that they hadn’t been produced.

13 96. Plaintiffs deem this to be a “compelling justification” but are open to alternate
14 approaches which would provide the same information. One alternative would be adding the
15 metadata field “Missing Google Drive Attachments” which would include the links to all linked
16 Google Drive documents which could not be retrieved. This field, and another added metadata
17 field “Non-Contemporaneous,” which would contain the value “Y” if a produced Google Drive
18 document was not the contemporaneous version, were both described by Plaintiffs in Plaintiffs
19 Proposed Methodology for Retrieving Google Drive Documents Linked to Within Google Emails
20 (Paragraphs 7, 8).

21 97. Additionally, Plaintiffs seek clarification of Uber’s ALLCUSTODIAN metadata
22 field which is defined in Appendix 2 of Uber’s proposed ESI protocol in relevant part as follows:
23 “Name(s) of person(s) or other data source (non-human) from where documents/files are
24 produced.” When there are multiple custodians of a document, Plaintiffs will not be able to
25 determine from which custodian or other data source the produced instance was collected. To
26 resolve this issue, Plaintiffs request clarification that the custodian or data source from whom the
27 produced version of a document was collected must be listed first in the ALLCUSTODIAN field.

1 98. The Meta Pixel Declarations and Order

2 99. In PTO No. 9, the Court, referring to the Parties’ description of programs such as
3 Metasploit’s *Forensic Email Collector* (“FEC”) and *Google Vault*, stated that in “other complex
4 litigation, more detailed information has been requested and provided to explain why such tools
5 are not feasible,” citing two declarations and an order from *In re: Meta Pixel Healthcare*
6 *Litigation* (No. 22-cv-03580-WHO (VKD))²² and quoting from the order’s ruling (“the
7 commercially available tools plaintiffs suggest may be used for automatically collecting links to
8 non-public documents have no or very limited utility in Meta’s data environments or systems”).

9 100. None of the complexities addressed by the Meta Pixel declarations and order are
10 present here.

11 101. Meta’s declarant Sam Yang, an internal Meta forensic resource, dismisses the use
12 of FEC, a tool for Google Mail environments that can collect Google Mail links to Google Drive
13 documents, in Meta’s environment, which uses Microsoft platforms along with Meta’s extensive
14 proprietary internally-developed platforms, and a non-Microsoft email archiving system. Yang
15 Decl. ¶¶ 10-11, 16-17, 20. I do not disagree with Mr. Yang’s conclusions in this regard.

16 102. Mr. Yang also notes that FEC does not collect linked documents from platforms
17 other than Google Workspace. Yang. Decl. ¶ 18. Again, I do not disagree with Mr. Yang’s
18 conclusions in this regard.

19 103. Mr. Yang does not claim experience with FEC or comment about its suitability and
20 capability when used, on live Google email and Google Drive documents.

21 104. The very reasons that Mr. Yang cites for FEC’s unsuitability in Meta’s
22 environment are the very reasons that FEC is appropriate for use here: unlike Meta, Uber *does* use
23 Google Mail which have links to Google Drive documents, the exact circumstances for which
24 FEC was designed.

25 ²² Declaration of Sam Yang, *In re: Meta Pixel Healthcare Litigation*, No. 22-cv-03580-WHO (VKD), Dkt. No. 265
(N.D. Cal. June 1, 2023); *id.*, Declaration of Jamie Brown, *In re: Meta Pixel Healthcare Litigation*, No. 22-cv-03580-
WHO (VKD), Dkt. No. 266 (N.D. Cal. June 1, 2023); Third Order re Dispute re ESI Protocol, *In re: Meta Pixel*
Healthcare Litigation, No. 22-cv-03580-WHO (VKD), Dkt. No. 267 (N.D. Cal. June 2, 2023).

1 105. Mr. Yang's declaration goes on at length about the complexities and intricacies of
2 Meta's data environment²³, see, e.g., ¶¶ 8, 10-11, 20, the incompatibility of Microsoft's collection
3 tool eDiscovery Purview (Premium) with Meta's unique environment and the disruptions which
4 would occur with its use by Meta, ¶¶ 20-21, and the absence of tools which work across unrelated
5 platforms, i.e., where the documents which contain links and the linked documents are in different
6 platforms.

7 106. Again, this is not the case here with respect to links to active Google Drive
8 documents in active Google emails.

9 107. Meta's declarant James Brown is a Senior Vice President at Lighthouse, Meta's
10 vendor in Meta Pixel and also Uber's vendor here. Mr. Brown's declaration focuses exclusively
11 on the technical challenges of developing and implementing a program to collate metadata, such
12 as senders, etc., from in-thread emails into consolidated metadata fields to be produced for the
13 "Last in Time" emails which contain the in-thread emails. There is no need or requirement for
14 such a program here.

15 I declare under penalty of perjury that the foregoing is true and correct to the best of my
16 knowledge.

17 DATED: April 12, 2024

/s/Douglas Forrest
Douglas Forrest
Senior VP, eDiscovery
International Litigation
Services, Inc.
dforrest@ilsteam.com

23 One such intricacy raised by Mr. Yang is Meta's use of multiple tools to shorten the URLs of linked documents.
Yang Decl. ¶ 11. Uber has not raised such an issue in the meet and confers held in response to this Court's directive or
in the Data Sources Letter.